Autoencoding star formation from cosmological parameters

*Deep*Thought α Summer Student Project

Benjamin Hertzsch [1], Michele Mancarella [2], Lucas Lombriser [2]

[1] Department of Physics, ETH Zurich. [2] Department of Theoretical Physics, University of Geneva

1 Introduction

The *Deep*Thought α Project is an interdisciplinary research collaboration exploring the likelihood of life in our cosmos. Treating the cosmic star formation rate (SFR) as a proxy [6], this summer project investigated a novel method of compressing multivariate data in order to arrive at a notion of **equivalence of universes with different cosmological parameters**.

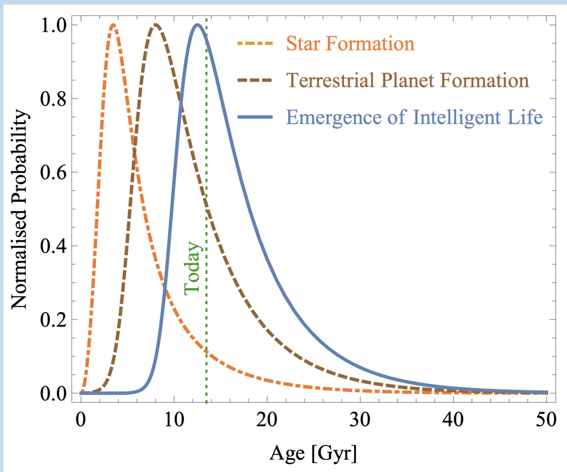

Figure 1: History of the normalized probabilities for star, planet and life formation in our cosmology [6]

2 Computational Method

Considering the statistical playground of a “multiverse”, a recent review [1] constrained the habitable ranges of numerous cosmological parameters (CPs). In this high-dimensional space, we are interested in the “eigenvectors” of **multiple-parameter variation**. We modify the concept of an autoencoder [3], encoding an input to a low-dimensional representation, $f(\boldsymbol{x}) \to \boldsymbol{z}$, and hence decoding an output, $g(\boldsymbol{z}) \to \boldsymbol{x}'$.

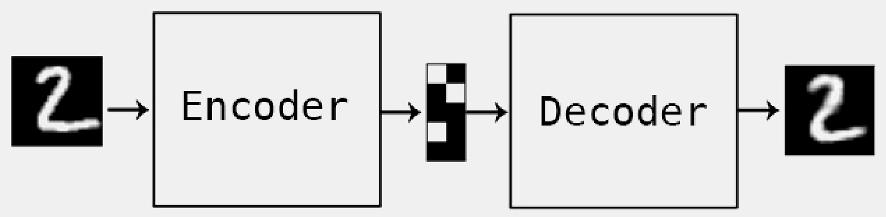

Figure 2: The concept of autoencoding, reproduced from https://blog.keras.io/building-autoencoders-in-keras.html

The concept was further developed to a generative network known as a **variational autoencoder** (VAE) [4] with the key features:

1. the encoder maps to a distribution in latent space
   $f(\boldsymbol{x}) \to \boldsymbol{\mu}, \boldsymbol{\sigma}^2$ s.t. $\boldsymbol{z} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\sigma}^2\right)$ and $g(\boldsymbol{z}) \to \boldsymbol{x}'$
2. the latent variables are stochastically sampled as
   $\boldsymbol{z} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \epsilon$ with $\epsilon_i \sim \mathcal{N}(0,1)$
3. learning is regularized by the KL-divergence associated to the latent distribution

Altogether, the input is compressed to a **low-dimensional disentangled latent representation**.

Treating the latent distribution as a multivariate diagonal Gaussian, one obtains the succinct loss function [4]

$$\mathcal{L} = \mathbb{E}\left[(\boldsymbol{x} - f(g(\boldsymbol{x})))^2\right] + \frac{\beta}{2}\sum_i \left(1 + \log\sigma_i^2 - \mu_i^2 - \sigma_i^2\right), \qquad (1)$$

where i ranges over the latent variables and $\beta \in \mathbb{R}_{>0}$ controls the disentanglement pressure. [2]

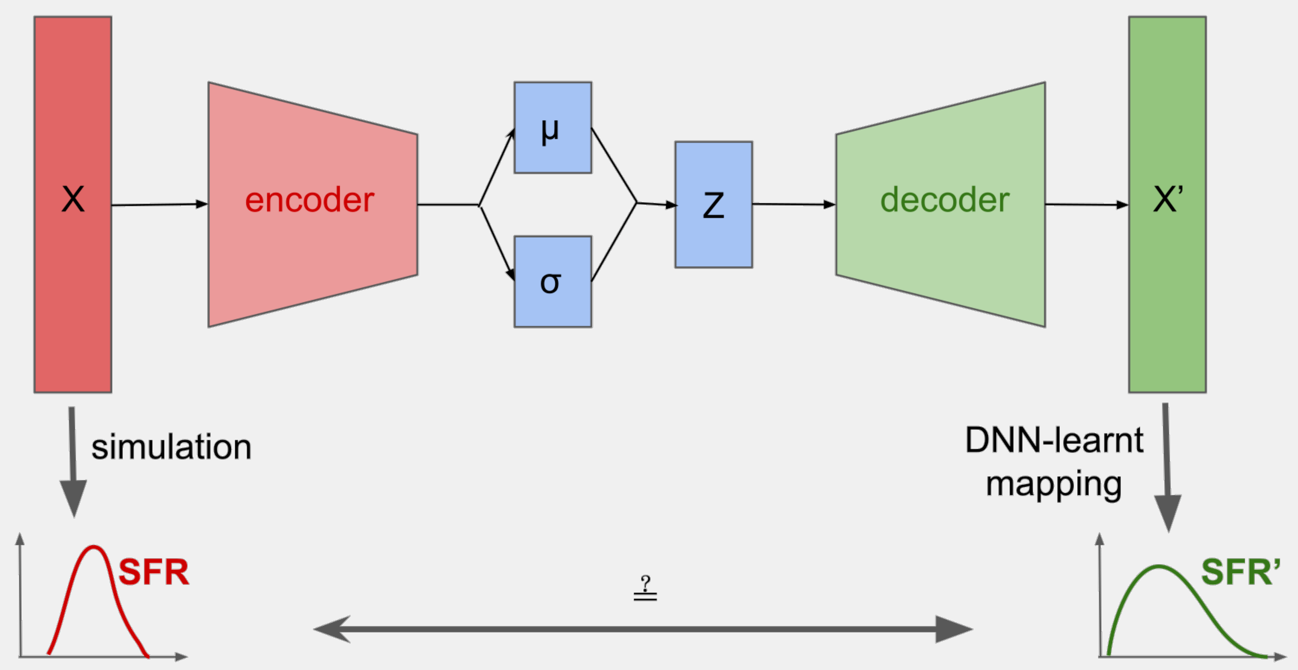

Figure 3: Scheme of the modified VAE-architecture developed and applied in this project

At this point we modify the architecture: instead of demanding equality of the reconstructed CPs, we **compare the corresponding SFRs**. Note that in order to circumvent time-consuming simulations in every training epoch, a second network is trained for mapping reconstructed CPs to SFRs; it is called within the VAE.

3 Data and Learning Results

While the project will eventually analyze multiversal SFR-simulations, the summer project aimed at developing the neural network architecture. For this purpose, we made use of an analytical toy model [5] fitting normalized SFRs dependent on 2 CPs,

$$\frac{\dot{\rho}_*}{\rho_{c0}} \propto \frac{f_{\text{in}}(z) + w(z)\cdot f_{\text{out}}(z)}{1 + w(z)} \qquad (2)$$

with $f_{\text{in}}(z)$, $f_{\text{out}}(z)$, $w(z)$ being functions of the redshift z, and we introduced an additional parameter controlling the SFR-amplitude.

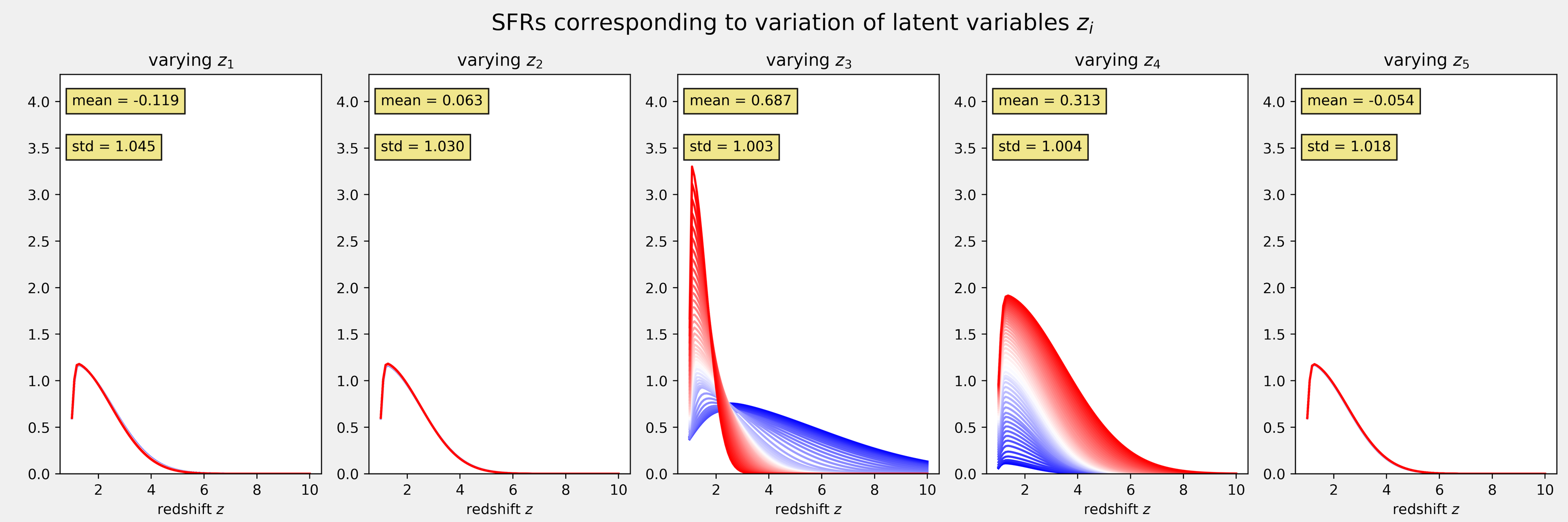

Figure 4: Having trained 7 input CPs to 5 latent variables, the SFRs obtained upon variation of the z_i are shown. The colour scale indicates regularly spaced $z_i \in \mu_i \pm 2\sigma_i$, fixing the other $z_j = \mu_j$ respectively

Applying this toy model, an exemplary compression of 7 CPs (of which accordingly 3 are relevant for the SFR) to a 5-dimensional latent representation with $\beta = 1$ was performed. Inspecting Fig. 4 and the correlation of latent variables over all datapoints, it is apparent that the network succeeded at learning disentangled features while “switching off” the remaining variables.

4 Conclusion

Having successfully built the modified VAE-architecture, it will be applied to multiversal SFR-simulations in future research. Here, the main challenges for learning analysis are:

1. mathematically formalizing the learned parameter equivalence
2. assigning statistical uncertainties

Equipped with these, one can explore new avenues for calculating the likelihood of life in the multiverse.

References

[1] F. C. Adams. The degree of fine-tuning in our universe – and others. *Physics Reports*, 2019.

[2] I. Higgins et al. Early visual concept learning with unsupervised deep learning, 2016.

[3] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science*, 313(5786):504–507, 2006.

[4] D. P. Kingma and M. Welling. Auto-encoding variational bayes, 2014.

[5] L. Lombriser and V. Smer-Barreto. Is there another coincidence problem at the reionization epoch? *Phys. Rev. D*, 96, Dec 2017.

[6] D. Sobral-Blanco and L. Lombriser. Exploring the self-tuning of the cosmological constant from planck mass variation, 2021.